

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-090997

(43)Date of publication of application : 04.04.1997

(51)Int.Cl.

G10L 9/14  
H03H 17/02  
H03H 17/02  
H03M 7/30

(21)Application number : 07-247827

(71)Applicant : MITSUBISHI ELECTRIC CORP

(22)Date of filing : 26.09.1995

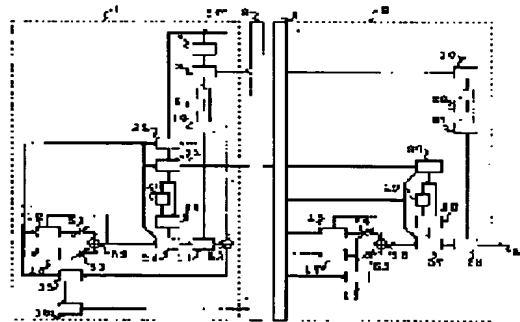
(72)Inventor : YAMAURA TADASHI  
TAKAHASHI MASAYA

## (54) SPEECH CODING DEVICE, SPEECH DECODING DEVICE, SPEECH CODING/DECODING METHOD AND COMPOSITE DIGITAL FILTER

### (57)Abstract:

PROBLEM TO BE SOLVED: To synthesize speech with high quality by making a degree of a filter variable which adds a short-term phase amplitude characteristic to a sound source signal in a speech coding device.

SOLUTION: A phase amplitude characteristic adding filter 34 of a coding part 1 varies a filter degree corresponding to the degree inputted from an optimal sound source searching means 35, multiplying an adaptive sound source vector of a sub-frame length outputted from an adaptive sound source sign table 12 and a driving sound vector of a sub-frame length outputted from a driving sound source sign table 13 by sound source gains  $\beta$  and  $\gamma$  in amplifiers 50 and 51 respectively, adding a phase-amplitude characteristic interpolated to the sound source signal to which each of the above resultants is added in an adder 52, outputting the results to a synthetic filter 11. Also, a phase-amplitude characteristic adding filter 37 of a decoding part 2 varies a filter degree corresponding to the degree of the phase-amplitude characteristic, multiplying the adaptive sound source vector and the driving sound source vector by individual sound source gains  $\beta$  and  $\gamma$  respectively, adding a phase-amplitude characteristic interpolated to the sound source signal obtained by adding each of the above resultants in an adder 56, outputting the results to a synthetic filter 12.



### LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

## (12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-90997

(43) 公開日 平成9年(1997)4月4日

(51) Int.Cl. <sup>8</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 9/14			G 1 0 L 9/14	J
				H
H 0 3 H 17/02	6 0 1	9274-5J	H 0 3 H 17/02	6 0 1 H
	6 3 5	9274-5J		6 3 5 C
H 0 3 M 7/30		9382-5K	H 0 3 M 7/30	Z
審査請求 未請求 請求項の数11 O L (全 18 頁)				

(21) 出願番号 特願平7-247827

(22) 出願日 平成7年(1995)9月26日

(71) 出願人 000006013

三菱電機株式会社

東京都千代田区丸の内二丁目2番3号

(72) 発明者 山浦 正

東京都千代田区丸の内二丁目2番3号 三

菱電機株式会社内

(72) 発明者 高橋 真哉

東京都千代田区丸の内二丁目2番3号 三

菱電機株式会社内

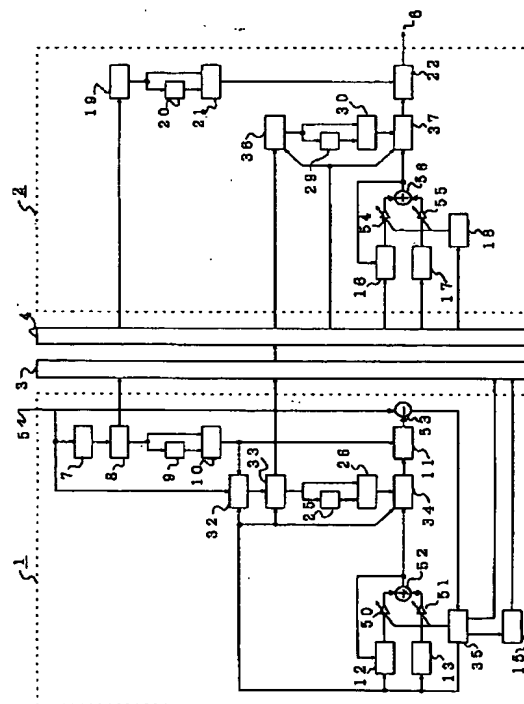
(74) 代理人 弁理士 宮田 金雄 (外3名)

(54) 【発明の名称】 音声符号化装置、音声復号化装置、音声符号化復号化方法および複合デジタルフィルタ

## (57) 【要約】

【課題】 従来装置は、位相振幅特性付加フィルタの次数は固定で、長いピッチ周期に対応するためには、フィルタ次数を大きくとる必要があるが、ピッチ周期が短い音声信号に対しては位相振幅特性付加フィルタのインパルス応答が複数ピッチに渡って影響を与え、合成音声の品質が劣化する。これを防ぎ、高品質音声を合成する。

【解決手段】 音源信号生成手段と、前記音源信号生成手段から出力される音源信号に位相振幅特性を付加する位相振幅特性付加フィルタと、前記位相振幅特性を付加した音源信号から合成音声を生成する合成フィルタとを備える符号駆動線形予測(CELP)符号化装置において、前記位相振幅特性付加フィルタはその次数を可変に構成する。



## 【特許請求の範囲】

【請求項1】 音源信号生成手段と、前記音源信号生成手段から出力される音源信号に位相振幅特性を付加する位相振幅特性付加フィルタと、前記位相振幅特性を付加した音源信号から合成音声を生成する合成フィルタとを備える符号駆動線形予測(CELP)符号化装置において、前記位相振幅特性付加フィルタはその次数が可変に構成されたことを特徴とする音声符号化装置。

【請求項2】 請求項1に記載の音声符号化装置において、前記位相振幅特性付加フィルタの次数は入力音声の特徴に応じて変更する構成にされたことを特徴とする音声符号化装置。

【請求項3】 請求項1に記載の音声符号化装置において、前記位相振幅特性付加フィルタの次数は入力音声を符号化した符号に応じて変更する構成にされたことを特徴とする音声符号化装置。

【請求項4】 音源信号生成手段と、前記音源信号生成手段から出力される音源信号に短期の位相振幅特性を付加する位相振幅特性付加フィルタと、前記位相振幅特性を付加した音源信号から合成音声を生成する合成フィルタとを備える符号駆動線形予測(CELP)復号化装置において、位相振幅特性付加フィルタはその次数が可変に構成されたことを特徴とする音声復号化装置。

【請求項5】 請求項4に記載の音声復号化装置において、前記位相振幅特性付加フィルタの次数は復号化した符号に応じて変更する構成にされたことを特徴とする音声復号化装置。

【請求項6】 符号化側においては、入力音声信号を線形予測分析して得られる線形予測パラメータを符号化し、入力音声信号の線形予測残差信号を分析して得られる短期の位相振幅特性を符号化し、音源符号帳より最適な合成音声を生成する音源信号を選択、符号化する一方、復号化側においては、入力された符号に基づき音源信号と位相振幅特性と線形予測パラメータとを生成し、出力音声信号を得る符号駆動線形予測(CELP)符号化復号化方法において、符号化側は、位相振幅特性付加フィルタの次数を変更する工程を含み、復号化側は、位相振幅特性付加フィルタの次数を前記符号化側の位相振幅特性付加フィルタの次数に対応して変更する工程を含むことを特徴とする音声符号化復号化方法。

【請求項7】 音源信号生成手段と、前記音源信号生成手段から出力される音源信号から合成音声を生成する合成フィルタと、入力音声と前記合成フィルタで生成された合成音声との歪みを評価し、歪みが最小となる音源信号を決定する最適音源探索手段とを備え、一定区間毎に音源信号情報を符号化する符号駆動線形予測(CELP)符号化装置において、前記合成フィルタは前記一定区間に加え、該区間以後所定の区間も予測して合成音声を生成する構成とし、前記

最適音源探索手段は前記合成音声生成区間に対応する入力音声と前記合成音声との歪みを評価し、歪みが最小となる音源信号を決定する構成としたことを特徴とする音声符号化装置。

【請求項8】 複数のデジタルフィルタから構成され、各フィルタのフィルタ特性は離散時間点で与えられる複合デジタルフィルタであって、各フィルタ特性を補間する際に、前記複合デジタルフィルタのフィルタ特性を補間した特性を得るフィルタ特性のフィルタを備えたことを特徴とする複合デジタルフィルタ。

【請求項9】 請求項8に記載の複合デジタルフィルタを用いることを特徴とする音声符号化装置または音声復号化装置。

【請求項10】 複数のデジタルフィルタから構成され、各フィルタのフィルタ係数は離散時間点で与えられる複合デジタルフィルタであって、各フィルタ特性を補間する際に、複合デジタルフィルタのフィルタ特性を補間した特性が得られるように各フィルタの特性を設定する特性補間設定手段を備えたことを特徴とする複合デジタルフィルタ。

【請求項11】 請求項10に記載の複合デジタルフィルタを用いることを特徴とする音声符号化装置または音声復号化装置。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】この発明は、音声信号をデジタル信号に圧縮符号化する符号駆動線形予測音声符号化装置と、前記圧縮符号を復号化する符号駆動線形予測音声復号化装置、符号化復号化方法、およびこれらに使用可能なデジタルフィルタ補間法に関する。

## 【0002】

【従来の技術】図9は、従来の符号駆動線形予測符号化復号化装置の全体構成の一例を示すものであり、W.B.Kleijn, D.J.Krasinski, R.H.Ketchum著 "Improved speech quality and efficient vector quantization in CELP" (ICASSP 88, pp.155-158, 1988)に示されたのと同様のものである。

【0003】図において、1は符号化部、2は復号化部、3は多重化手段、4は分離手段であり、5は入力音声、6は出力音声である。7は線形予測パラメータ分析手段、8は線形予測パラメータ符号化手段、9、20は遅延器、10、21は線形予測パラメータ補間手段、11、22は合成フィルタである。12、16は適応音源符号帳、13、17は駆動音源符号帳であり、14は最適音源探索手段、15は音源利得符号化手段である。18は音源利得復号化手段、19は線形予測パラメータ復号化手段である。50、51、54、55は増幅器、52、56は加算器、53は減算器である。

【0004】以下、上記従来の符号駆動線形予測符号化

復号化装置の動作について説明する。

【0005】まず符号化部1において、線形予測パラメータ分析手段7は、入力音声5を一定のフレーム周期で分析して、線形予測パラメータを抽出する。次いで線形予測パラメータ符号化手段8が前記線形予測パラメータを量子化し、それに対応する符号を多重化手段3に出力すると共に、量子化した線形予測パラメータを遅延器9と線形予測パラメータ補間手段10に出力する。遅延器9は前記量子化した線形予測パラメータを1フレーム分時間遅延して線形予測パラメータ補間手段10に出力する。すなわち遅延器9は前フレームの量子化した線形予測パラメータを線形予測パラメータ補間手段10に出力する。線形予測パラメータ補間手段10は前記遅延器9から入力される前フレームの量子化した線形予測パラメータと前記線形予測パラメータ符号化手段8から入力される現フレームの量子化した線形予測パラメータを図1

$$H_n(z) = \frac{1}{1 + \sum_{k=1}^{N_h} a_n(k)z^{-k}} \quad (1)$$

【0008】

$$a_i(k) = w_i a_{i-1}(k) + (1 - w_i) a_n(k) \quad (k = 1, \dots, N_h) \quad (2)$$

【0009】次に、音源信号情報の符号化について説明する。音源信号情報は例えば1/4フレームを単位とするサブフレーム毎に符号化を行なう。音源信号情報を符号化するサブフレームと線形予測パラメータを分析、補間する時間関係の例は図10に示されるものと同様であり、線形予測パラメータはフレームの最後のサブフレームを対象に分析され、それ以外のサブフレームの線形予測パラメータは現フレームと前フレームで分析された線形予測パラメータを補間することで求められる。

【0010】適応音源符号帳12には過去に生成した音源信号が記憶されており、最適音源探索手段14より入力される適応音源符号Lに対応した適応音源ベクトルを出力する。駆動音源符号帳13には、例えばランダム雑音から生成したn個の駆動音源ベクトルが記憶されており、最適音源探索手段14より入力される駆動音源符号Iに対応した駆動音源ベクトルを出力する。前記適応音源ベクトル及び駆動音源ベクトルのベクトル長はサブフレーム長と一致している。合成フィルタ11は、前記適応音源ベクトル及び前記駆動音源ベクトルにそれぞれ音源利得 $\beta$ 、 $\gamma$ を増幅器50、51で乗じ、これ等を加算器53で加算した音源信号と、前記補間した線形予測パラメータとを用いてサブフレーム長の合成音声を生成する。

【0011】最適音源探索手段14は、前記合成音声と入力音声5との該サブフレーム間における減算器53で算出された誤差信号の聴覚重み付き歪みを評価し、前記歪みが最小になる適応音源符号L、駆動音源符号I、

0に示す方法で、例えば1/4フレーム周期毎に補間し、補間した線形予測パラメータを合成フィルタ11に出力する。図10は線形予測パラメータを分析、補間する時間関係の例を示し、線形予測パラメータはフレームの最後のサブフレームを対象に分析され、それ以外のサブフレームの線形予測パラメータは現フレームと前フレームで分析された線形予測パラメータを補間することで求められる。

【0006】ここで、合成フィルタ11は例えばフレームnで求められた伝達関数 $H_n(z)$ が式(1)で表現される $N_h$ 次のフィルタであり、線形予測パラメータ補間手段10は例えば式(2)に従ってフレームn-1とフレームnとで求められたフィルタ係数 $a_{n-1}(k)$ 、 $a_n(k)$ を線形に補間してフィルタ係数 $a_i(k)$ を求める。

【0007】

【数1】

【数2】

音源利得 $\beta$ 、 $\gamma$ を求め、適応音源符号Lと駆動音源符号Iを多重化手段3に出力するとともに、音源利得 $\beta$ 、 $\gamma$ を音源利得符号化手段15に出力する。音源利得符号化手段15は、前記音源利得 $\beta$ 、 $\gamma$ を量子化し、その符号を多重化手段3に出力する。

【0012】上記の適応音源符号帳12は、前記歪みが最小になる適応音源符号Lに対応する適応音源ベクトル、駆動音源符号Iに対応する駆動音源ベクトル及び量子化した音源利得 $\beta$ 、 $\gamma$ を用いて生成した音源信号によって符号帳の内容を更新する。

【0013】以上符号化が終了した後、多重化手段3は前記量子化した線形予測パラメータに対応する符号、適応音源符号L、駆動音源符号I、及び量子化した音源利得 $\beta$ 、 $\gamma$ に対応する符号を伝送路に送出する。

【0014】次に、復号化部2の動作について説明する。

【0015】まず多重化手段3の出力を受けた分離手段4は、その出力を分離し、

適応音源符号L→適応音源符号帳16

駆動音源符号I→駆動音源符号帳17

音源利得の符号→音源利得復号化手段18

線形予測パラメータの符号→線形予測パラメータ復号化手段19

にそれぞれ出力する。

【0016】適応音源符号帳16は前記適応音源符号Lに対応した適応音源ベクトルを出力し、駆動音源符号帳17は前記駆動音源符号Iに対応した駆動音源ベクトル

を出力する。

【0017】また、音源利得復号化手段18は前記音源利得の符号に対応した音源利得 $\beta$ 、 $\gamma$ を復号化し、前記適応音源ベクトル及び前記駆動音源ベクトルにそれぞれ音源利得 $\beta$ 、 $\gamma$ を乗じるように増幅器54、55を制御する。

【0018】一方、線形予測パラメータ復号化手段19は、前記線形予測パラメータの符号に対応する線形予測パラメータを復号化し、遅延器20と線形予測パラメータ補間手段21に出力する。遅延器20は前記復号化した線形予測パラメータを1フレーム分時間遅延して線形予測パラメータ補間手段21に出力する。すなわち遅延器20は前フレームの復号化した線形予測パラメータを線形予測パラメータ補間手段21に出力する。線形予測パラメータ補間手段21は前記遅延器20から入力される前フレームの復号化した線形予測パラメータと前記線形予測パラメータ復号化手段19から入力される現フレームの復号化した線形予測パラメータを図10に示すように、例えば1/4フレーム周期毎に補間し、補間した線形予測パラメータを合成フィルタ22に出力する。合成フィルタ22は、前記適応音源ベクトル及び前記駆動音源ベクトルを加算器56で加算して得られる音源信号を、前記線形予測パラメータを用いて合成し、出力音声6を出力する。

【0019】上記の適応音源符号帳16は、符号化部1の適応音源符号帳12と同様に、前記音源信号で符号帳の内容を更新する。

【0020】以上の従来例とは別に、他の符号化復号化装置として図11に示すものがある。

【0021】図11は、山浦、高橋著「低ビットレート CELPにおける有声音品質改善の検討」(日本音響学会平成6年度秋季研究発表会講演論文集、pp. 263-264、1994)に示されたのと同様のもので、音源信号の位相振幅特性を符号化するものである。

【0022】図11において図9と同一の部分については同一の符号を付し、その説明は省略する。

$$P_n(z) = \sum_{k=0}^{N_p-1} h_n(k) z^{-k} \quad (3)$$

【0028】

$$h_i(k) = w_i h_{i-1}(k) + (1 - w_i) h_n(k) \quad (k = 0, \dots, N_p - 1) \quad (4)$$

【0029】位相振幅特性付加フィルタ27は例えば位相振幅特性の分析次数のFIRフィルタであり、適応音源符号帳12から出力されるサブフレーム長の適応音源ベクトル及び駆動音源符号帳13から出力されるサブフレーム長の駆動音源ベクトルにそれぞれ音源利得 $\beta$ 、 $\gamma$ を増幅器50、51で乗じ、これ等を加算器53で加算した音源信号に対して、前記補間した位相振幅特性を付加し、合成フィルタ11に出力する。合成フィルタ11は、線形予測パラメータ補間手段10より入力される補

【0023】図9と異なる部分として、23は位相振幅特性分析手段、24は位相振幅特性符号化手段、25、29は遅延器、26、30は位相振幅特性補間手段、27、31位相振幅特性付加フィルタ、28は位相振幅特性復号化手段である。

【0024】ここで上記の構成による符号化復号化装置の動作について説明する。

【0025】まず、符号化部1において、位相振幅特性分析手段23は、入力音声5と線形予測パラメータ補間手段10より入力される線形予測パラメータを用いて線形予測残差信号を生成し、前記線形予測残差信号をフレーム周期で分析して短期の位相振幅特性を求め、位相振幅特性符号化手段24に出力する。このとき、位相振幅特性の分析次数は、例えば128次とする。位相振幅特性符号化手段24は、前記位相振幅特性を例えばベクトル量子化するなどして量子化し、それに対応する符号を多重化手段3に出力するとともに、量子化した位相振幅特性を遅延器25と位相振幅特性補間手段26に出力する。遅延器25は前記量子化した位相振幅特性を1フレーム分時間遅延して位相振幅特性補間手段26に出力する。すなわち、遅延器25は前フレームの量子化した位相振幅特性を位相振幅特性補間手段26に出力する。位相振幅特性補間手段26は前記遅延器25から入力される前フレームの量子化した位相振幅特性と前記位相振幅特性符号化手段24から入力される現フレームの量子化した位相振幅特性を図10に示すように、例えば1/4フレーム周期毎に補間し、補間した位相振幅特性を位相振幅特性付加フィルタ27に出力する。

【0026】ここで、位相振幅特性付加フィルタ27は例えばフレーム $n$ で求められた伝達関数 $P_n(z)$ が式(3)で表現される $N_p$ 次のフィルタであり、位相振幅特性補間手段26は例えば式(4)に従ってフレーム $n-1$ とフレーム $n$ とで求められたフィルタ係数 $h_{n-1}(k)$ 、 $h_n(k)$ を線形に補間してフィルタ係数 $h_i(k)$ を求める。

【0027】

【数3】

【数4】

間した線形予測パラメータと前記位相振幅特性を付加した音源信号とを用いてサブフレーム長の合成音声を生成する。

【0030】音源信号に位相振幅特性を付加して合成音声を生成する過程の説明図を図12に示す。

【0031】最適音源探索手段14は、前記合成音声と入力音声5との該サブフレーム区間における減算器53で算出された誤差信号の聴覚重み付き歪みを評価し、前記歪みが最小になる適応音源符号 $L$ 、駆動音源符号 $I$ 、

音源利得 $\beta$ 、 $\gamma$ を求め、適応音源符号 $L$ と駆動音源符号 $I$ を多重化手段3に出力するとともに、音源利得 $\beta$ 、 $\gamma$ を音源利得符号化手段15に出力する。音源利得符号化手段15は、前記音源利得 $\beta$ 、 $\gamma$ を量子化し、その符号を多重化手段3に出力する。

【0032】以上符号化が終了した後、多重化手段3は前記量子化した線形予測パラメータに対応する符号、量子化した位相振幅特性に対応する符号、適応音源符号 $L$ 、駆動音源符号 $I$ 、及び量子化した音源利得 $\beta$ 、 $\gamma$ に対応する符号を伝送路に送出する。

【0033】次に、復号化部2の動作について説明する。

【0034】まず、多重化手段3の出力を受けた分離手段4は、その出力を分離し、

適応音源符号 $L$ →適応音源符号帳16

駆動音源符号 $I$ →駆動音源符号帳17

音源利得の符号→音源利得復号化手段18

位相振幅特性の符号→位相振幅特性復号化手段28

線形予測パラメータの符号→線形予測パラメータ復号化手段19

にそれぞれ出力する。

【0035】位相振幅特性復号化手段28は、前記位相振幅特性の符号に対応する位相振幅特性を復号化し、遅延器29と位相振幅特性補間手段30に出力する。遅延器29は前記復号化した位相振幅特性を1フレーム分遅延して位相振幅特性補間手段30に出力する。すなわち遅延器29は前フレームの復号化した位相振幅特性を位相振幅特性補間手段30に出力する。位相振幅特性補間手段30は前記遅延器29から入力される前フレームの復号化した位相振幅特性と前記位相振幅特性復号化手段28から入力される現フレームの復号化した位相振幅特性を図10に示すように、符号化部1の位相振幅特性補間手段26と同様、例えば1/4フレーム周期毎に補間し、補間した位相振幅特性を位相振幅特性付加フィルタ31に出力する。

【0036】位相振幅特性付加フィルタ31は、適応音源符号帳16から出力される適応音源ベクトル及び駆動音源符号帳17から出力される駆動音源ベクトルにそれぞれ音源利得復号化手段18から出力される音源利得 $\beta$ 、 $\gamma$ を増幅器54、55で乗じ、これ等を加算器56で加算して得られる音源信号に対して、前記復号化した位相振幅特性を付加し、合成フィルタ22に出力する。合成フィルタ22は、前記位相振幅特性を付加した音源信号を、線形予測パラメータ補間手段21より入力される線形予測パラメータを用いて合成し、出力音声6を出力する。

【0037】

【発明が解決しようとする課題】音声には有声音と無声音とがあり、有声音の再現性が合成音声の品質に与える影響は大きい。ここでこの有声音の音源は、ピッチ周期

性とピッチ周期における短期の位相特性をもつ信号としてモデル化することができる。そして、このピッチ周期は話者により2～18msの値をとり、男声では7.5～10ms程度、女声では2.5～5ms程度に多く分布している。

【0038】上記した従来の符号駆動線形予測符号化・復号化装置では、音源信号を適応音源ベクトルと駆動音源ベクトルとの加算で表すが、この方法は音源信号の位相特性を直接的に表現するものではない。従って、音源信号の位相特性を再現できない場合が生じ、合成音声の品質が劣化するという問題があった。

【0039】これに対して、従来の音源信号の位相振幅特性を符号化する符号駆動線形予測符号化復号化装置では、音源信号に位相振幅特性を付加するフィルタを備えることにより、音源信号の位相特性の再現性の向上を図っている。しかし、位相振幅特性付加フィルタの次数は固定であり、長いピッチ周期に対応するためにフィルタ次数を大きくとる必要があるが、このときピッチ周期が短い音声信号に対しては位相振幅特性付加フィルタのインパルス応答が複数ピッチに渡って影響を与えることになり、ピッチ周期における短期の位相特性を与えるという目的から外れ、合成音声の品質が劣化するという問題があった。

【0040】図13はピッチ周期と位相振幅特性付加フィルタのインパルス応答長の関係を示すものであり、図13(a)はピッチ周期とインパルス応答長が同程度の場合、図13(b)はピッチ周期がインパルス応答長よりも短い場合であり、この図13(b)からも明かなように位相振幅特性付加フィルタのインパルス応答が複数ピッチに渡って影響を与えることが分かる。

【0041】また、従来の符号駆動線形予測符号化復号化装置及び従来の音源信号の位相振幅特性を符号化する符号駆動線形予測符号化復号化装置では、音源符号帳より最適な音源信号を選択する際に、音源信号を決定するサブフレーム区間における合成音声と入力音声との歪みを評価して音源信号を決定していた。

【0042】しかし、該サブフレーム区間の音源信号は、合成フィルタや位相振幅特性付加フィルタのフィルタ応答により、該区間以後の合成音声にも影響を与えるので、該サブフレーム区間で歪みが小さくても、該区間以後で大きな歪みを生じる場合があり、合成音声の品質が劣化するという問題があった。

【0043】さらに、従来の音源信号の位相振幅特性を符号化する符号駆動線形予測符号化復号化装置では、位相振幅特性付加フィルタのフィルタ特性と合成フィルタのフィルタ特性を別々に補間しているが、その補間特性はこれら2つのフィルタを複合した1つのフィルタとして補間した場合の補間特性とは異なっており、合成音声の品質が劣化するという問題があった。

【0044】この発明は、かかる課題を解決するために

なされたもので、その目的は、音声を符号化復号化するに当たり、合成音声の品質の劣化を回避し、品質の良い合成音声を生成することができる符号駆動線形予測符号化復号化装置及び方法を得ることにある。

【0045】

【課題を解決するための手段】上記の課題を解決するために、この発明の音声符号化装置は、音源信号生成手段と、前記音源信号生成手段から出力される音源信号に短期の位相振幅特性を付加する位相振幅特性付加フィルタと、前記位相振幅特性を付加した音源信号から合成音声を生成する合成フィルタとを備える符号駆動線形予測(CELP)符号化装置において、位相振幅特性付加フィルタの次数を可変とするものである。

【0046】また、この発明の音声符号化装置はさらに、入力音声の特徴に応じて位相振幅特性付加フィルタの次数を変更するものである。

【0047】また、この発明の音声符号化装置は、入力音声を符号化した符号に応じて位相振幅特性付加フィルタの次数を変更するものである。

【0048】一方、この発明の音声復号化装置は、音源信号生成手段と、前記音源信号生成手段から出力される音源信号に短期の位相振幅特性を付加する位相振幅特性付加フィルタと、前記位相振幅特性を付加した音源信号から合成音声を生成する合成フィルタとを備える符号駆動線形予測(CELP)復号化装置において、位相振幅特性付加フィルタの次数を可変とするものである。

【0049】また、この発明の音声復号化装置はさらに、復号化した符号に応じて位相振幅特性付加フィルタの次数を変更するものである。

【0050】また、この発明に係る音声符号化復号化方法は、符号化側において、入力音声信号を線形予測分析して得られる線形予測パラメータを符号化し、入力音声信号の線形予測残差信号を分析して得られる短期の位相振幅特性を符号化し、音源符号帳より最適な合成音声を生成する音源信号を選択、符号化する一方、復号化側において、入力された符号に基づき音源信号と位相振幅特性と線形予測パラメータとを生成し、出力音声信号を得る符号駆動線形予測(CELP)符号化復号化方法であって、符号化側は、位相振幅特性付加フィルタの次数を変更する工程を付加し、復号化側は、位相振幅特性付加フィルタの次数を前記符号化側の位相振幅特性付加フィルタの次数に対応して変更する工程を付加するものである。

【0051】また、この発明の音声符号化装置は、音源信号生成手段と、前記音源信号生成手段から出力される音源信号から合成音声を生成する合成フィルタと、入力音声と前記合成フィルタで生成された合成音声との歪みを評価し、歪みが最小となる音源信号を決定する最適音源探索手段とを備え、一定区間毎に音源信号情報を符号化する符号駆動線形予測(CELP)符号化装置において、音源信号情報を符号化する際に、前記合成フィルタは前記

一定区間に加え、該区間以後において生成される合成音声を予測して生成し、最適音源探索手段は前記合成音声生成区間に対応する入力音声と前記合成音声との歪みを評価し、音源信号を決定するものである。

【0052】また、この発明の複合デジタルフィルタは、複数個のデジタルフィルタから構成され、各フィルタのフィルタ特性は離散時間点で与えられる複合デジタルフィルタであって、各フィルタ特性を補間する際、前記複合デジタルフィルタのフィルタ特性を補間した特性を得るフィルタ特性のフィルタを備えたものである。

【0053】また、この発明の音声符号化装置または音声復号化装置は、複数個のデジタルフィルタをその構成要素に持ち、各フィルタのフィルタ特性は離散時間点で与えられるものにおいて、各フィルタ特性を補間する際、前記複合デジタルフィルタのフィルタ特性を補間した特性を得るフィルタ特性のフィルタを備えたものである。

【0054】また、この発明の複合デジタルフィルタは、複数個のデジタルフィルタから構成され、各フィルタのフィルタ係数は離散時間点で与えられる複合デジタルフィルタであって、各フィルタ特性を補間する際、複合デジタルフィルタのフィルタ特性を補間した特性が得られるように各フィルタの特性を設定する特性補間設定手段を備えたものである。

【0055】また、この発明の音声符号化装置または音声復号化装置は、複数個のデジタルフィルタをその構成要素に持ち、各フィルタのフィルタ特性は離散時間点で与えられるものにおいて、各フィルタ特性を補間する際、複合デジタルフィルタのフィルタ特性を補間した特性が得られるように各フィルタの特性を設定する特性補間設定手段を備えたものである。

【0056】

【発明の実施の形態】

実施の形態1. ここでこの発明に係る音声符号化装置及び音声復号化装置を図面に従って説明する。

【0057】図1は、本実施例の音声符号化装置および音声復号化装置の全体構成を示すブロック図である。この図において図11と同一の部分については同一の符号を付し、説明を省略する。

【0058】この実施の形態において新たな構成は、求める位相振幅特性の分析次数を変更して分析可能な位相振幅特性分析手段32、位相振幅特性の分析次数に対応して量子化する位相振幅特性符号化手段33、位相振幅特性の分析次数に対応してフィルタ次数を変更可能な位相振幅特性付加フィルタ34、37、最適な位相振幅特性付加フィルタの次数と音源符号の組合せを探索する最適音源探索手段35、及び符号化された位相振幅特性の次数に対応して復号化する位相振幅特性復号化手段36である。

【0059】以下、上記の新たな構成を中心に動作を説明する。

【0060】まず、符号化部1において、位相振幅特性分析手段32は、入力音声5と線形予測パラメータ補間手段10より入力される線形予測パラメータを用いて線形予測残差信号を生成し、フレーム毎に前記線形予測残差信号を最適音源探索手段35より入力される次数で短期の位相振幅特性を分析して求め、位相振幅特性符号化手段33に出力する。位相振幅特性符号化手段33は最適音源探索手段35より入力される次数に対応して、例えばベクトル量子化する際の量子化符号帳を前記次数により切替えるなどして前記位相振幅特性を量子化し、それに対応する符号を多重化手段3に出力するとともに、量子化した位相振幅特性を遅延器25と位相振幅特性補間手段26に出力する。図2(a)に従来の位相振幅特性符号化手段24、(b)にこの実施の形態における位相振幅特性符号化手段33の構成の一例を示す。

【0061】図2(b)に示すように、位相振幅特性符号化手段33は複数の例えばM個の量子化符号帳を備え、切り替え手段により最適音源探索手段35より入力される次数に対応した量子化符号帳を量子化器に接続する。

【0062】遅延器25は前記量子化した位相振幅特性を1フレーム分時間遅延して位相振幅特性補間手段26に出力する。すなわち遅延器25は前フレームの量子化した位相振幅特性を位相振幅特性補間手段26に出力する。位相振幅特性補間手段26は前記遅延器25から入力される前フレームの量子化した位相振幅特性と前記位相振幅特性符号化手段33から入力される現フレームの量子化した位相振幅特性を図10に示す方法で、例えば1/4フレーム周期毎に補間し、補間した位相振幅特性を位相振幅特性付加フィルタ34に出力する。

【0063】位相振幅特性付加フィルタ34は例えばFIRフィルタであり、最適音源探索手段35から入力される次数に対応してフィルタ次数を変更し、適応音源符号帳12から出力されるサブフレーム（例えば1/4フレーム）長の適応音源ベクトル及び駆動音源符号帳13から出力されるサブフレーム長の駆動音源ベクトルにそれぞれ音源利得 $\beta$ 、 $\gamma$ を増幅器50、51で乗じ、これ等を加算器52で加算した音源信号に対して、前記補間した位相振幅特性を付加し、合成フィルタ11に出力する。合成フィルタ11は、線形予測パラメータ補間手段10より入力される補間した線形予測パラメータと前記位相振幅特性を付加した音源信号とを用いてサブフレーム長の合成音声を生成する。

【0064】最適音源探索手段35は、前記合成音声と入力音声5との該サブフレーム区間における減算器53で算出した誤差信号の聴覚重み付き歪みを評価し、前記歪みが最小になる位相振幅特性の次数、適応音源符号L、駆動音源符号I、音源利得 $\beta$ 、 $\gamma$ を求め、位相振幅特性の次数と適応音源符号Lと駆動音源符号Iとを多重

化手段3に出力するとともに、音源利得 $\beta$ 、 $\gamma$ を音源利得符号化手段15に出力する。音源利得符号化手段15は、前記音源利得 $\beta$ 、 $\gamma$ を量子化し、その符号を多重化手段3に出力する。

【0065】以上符号化が終了した後、多重化手段3は前記量子化した線形予測パラメータに対応する符号、量子化した位相振幅特性に対応する符号、位相振幅特性の次数、適応音源符号L、駆動音源符号I、及び量子化した音源利得 $\beta$ 、 $\gamma$ に対応する符号を伝送路に送出する。

【0066】以上がこの実施の形態の音声符号化装置に特徴的な動作である。

【0067】つづいて、復号化部2について説明する。

【0068】まず、多重化手段3の出力を受けた分離手段4は、その出力を分離し、

適応音源符号L→適応音源符号帳16

駆動音源符号I→駆動音源符号帳17

音源利得の符号→音源利得復号化手段18

位相振幅特性の符号→位相振幅特性復号化手段36

位相振幅特性の次数→位相振幅特性復号化手段36、位

相振幅特性付加フィルタ37

線形予測パラメータの符号→線形予測パラメータ復号化手段19

にそれぞれ出力する。

【0069】位相振幅特性復号化手段36は、前記位相振幅特性の次数に対応して、例えばベクトル量子化符号帳を前記次数により切替えるなどして、前記位相振幅特性の符号に対応する位相振幅特性を復号化し、遅延器29と位相振幅特性補間手段30に出力する。遅延器29は前記復号化した位相振幅特性を1フレーム分遅延して位相振幅特性補間手段30に出力する。すなわち遅延器29は前フレームの復号化した位相振幅特性を位相振幅特性補間手段30に出力する。位相振幅特性補間手段30は前記遅延器29から入力される前フレームの復号化した位相振幅特性と前記位相振幅特性復号化手段36から入力される現フレームの復号化した位相振幅特性とを符号化部1の位相振幅特性補間手段26と同様に例えば1/4フレーム周期毎に補間し、補間した位相振幅特性を位相振幅特性付加フィルタ37に出力する。

【0070】位相振幅特性付加フィルタ37は、前記位相振幅特性の次数に対応してフィルタ次数を変更し、適応音源符号帳16から出力される適応音源ベクトル及び駆動音源符号帳17から出力される駆動音源ベクトルにそれぞれ音源利得復号化手段18から出力される音源利得 $\beta$ 、 $\gamma$ を増幅器54、55で乗じ、これ等を加算器56で加算して得られる音源信号に対して、前記補間した位相振幅特性を付加し、合成フィルタ22に出力する。合成フィルタ22は、線形予測パラメータ補間手段21より入力される線形予測パラメータと前記位相振幅特性を付加した音源信号とを用いて出力音声6を合成し、出力する。



【0071】以上がこの実施の形態1の音声復号化装置に特徴的な動作である。

【0072】この実施の形態1によれば、音源信号に短期の位相振幅特性を付加する位相振幅特性付加フィルタのフィルタ次数を可変とし、入力音声と合成音声の歪みが小さくなるようにそのフィルタ次数を変更することにより、音源信号の再現性を良好にし、合成音声の品質を向上させることができる。

【0073】なお、この実施の形態1では位相振幅特性の符号化を、その次数により量子化符号帳を切替えているが、1つの量子化符号帳内に次数の異なる位相振幅特性を格納しておき、最適音源探索手段35は前記位相振幅特性の符号を指定し、位相振幅特性分析手段32及び位相振幅特性付加フィルタ34は前記位相振幅特性の符号に対応する次数により分析、フィルタリングを行なうとしてもよい。この場合、位相振幅特性の次数毎に異なる数の位相振幅特性を量子化符号帳内に格納しておくことができる。これにより、必要度が高い次数の位相振幅特性は多く、また、必要度が低い位相振幅特性は少なくすることができ、量子化効率を向上させることができる。

#### 【0074】実施の形態2

つづいてこの発明の音声符号化装置の実施の形態2を図面に従って説明する。

【0075】図3は、この実施の形態2の音声符号化装置の構成を示すブロック図である。この図において図1と同一の部分については同一の符号を付し、説明を省略する。

【0076】図3において、図1と比べて新たな構成は、入力音声のピッチ周期を分析して抽出するためのピッチ抽出手段38、ピッチ周期から位相振幅付特性の次数を決定する位相振幅特性次数決定手段39である。

【0077】以下、上記の追加された構成を中心に動作を説明する。

【0078】まず符号化部1において、ピッチ抽出手段38は既知の方法によって入力音声5のピッチ周期を分析して抽出し、位相振幅特性次数決定手段39に出力する。位相振幅特性次数決定手段39は、前記ピッチ周期に対応して位相振幅特性の次数を決定し、位相振幅特性分析手段32、位相振幅特性符号化手段33、位相振幅特性付加フィルタ34及び多重化手段3に出力する。前記位相振幅特性次数決定手段39においてピッチ周期に対する位相振幅特性の次数は、例えばピッチ周期が2～6msの時は40次、6～12msの時は80次、12ms以上では128次とする。

【0079】位相振幅特性分析手段32は、入力音声5と線形予測パラメータ補間手段10より入力される線形予測パラメータを用いて線形予測残差信号を生成し、フレーム毎に前記線形予測残差信号を前記位相振幅特性次数決定手段39より入力される次数で短期の位相振幅特

性を分析して求め、位相振幅特性符号化手段33に出力する。位相振幅特性符号化手段33は前記次数に対応して、例えばベクトル量子化する際の量子化符号帳を前記次数により切替えるなどして、前記位相振幅特性を量子化し、それに対応する符号を多重化手段3に出力するとともに、量子化した位相振幅特性を遅延器25と位相振幅特性補間手段26に出力する。

【0080】遅延器25は前記量子化した位相振幅特性を1フレーム分時間遅延して位相振幅特性補間手段26に出力する。すなわち遅延器25は前フレームの量子化した位相振幅特性を位相振幅特性補間手段26に出力する。位相振幅特性補間手段26は前記遅延器25から入力される前フレームの量子化した位相振幅特性の例えば1/4フレーム分と前記位相振幅特性符号化手段33から入力される現フレームの量子化した位相振幅特性の例えば1/4フレーム分とにより、この間を1/4フレーム（サブフレーム）周期毎に補間し、補間した位相振幅特性を位相振幅特性付加フィルタ34に出力する。

【0081】位相振幅特性付加フィルタ34は例えばFIRフィルタであり、前記位相振幅特性次数決定手段39から入力される次数に対応してフィルタ次数を変更し、適応音源符号帳12から出力されるサブフレーム長の適応音源ベクトル及び駆動音源符号帳13から出力されるサブフレーム長の駆動音源ベクトルにそれぞれ音源利得 $\beta$ 、 $\gamma$ を増幅器50、51で乗じ、これ等を加算器53で加算した音源信号に対して、前記補間した位相振幅特性を付加し、合成フィルタ11に出力する。合成フィルタ11は、線形予測パラメータ補間手段10より入力される補間した線形予測パラメータと前記位相振幅特性を付加した音源信号とを用いてサブフレーム長の合成音声を生成する。

【0082】最適音源探索手段14は、前記合成音声と入力音声5との該サブフレーム区間における減算器53で算出した誤差信号の聴覚重み付き歪みを評価し、前記歪みが最小になる適応音源符号L、駆動音源符号I、音源利得 $\beta$ 、 $\gamma$ を求め、適応音源符号Lと駆動音源符号Iとを多重化手段3に出力するとともに、音源利得 $\beta$ 、 $\gamma$ を音源利得符号化手段15に出力する。音源利得符号化手段15は、前記音源利得 $\beta$ 、 $\gamma$ を量子化し、その符号を多重化手段3に出力する。

【0083】以上符号化が終了した後、多重化手段3は前記量子化した線形予測パラメータに対応する符号、量子化した位相振幅特性に対応する符号、位相振幅特性の次数、適応音源符号L、駆動音源符号I、及び量子化した音源利得 $\beta$ 、 $\gamma$ に対応する符号を伝送路に送出する。

【0084】以上が実施の形態2に係る音声符号化装置の概要である。

【0085】この実施の形態によれば、入力音声进行分析して求めたピッチ周期に応じて位相振幅特性付加フィルタの次数を変更することにより、少ない処理量でも音源

信号の再現性を良好にし、合成音声の品質を向上させることができる。

#### 【0086】実施の形態3

さらにこの発明に係る音声符号化装置及び音声復号化装置の別の実施の形態を図面にしたがって説明する。

【0087】図4は、この実施の形態の音声符号化装置及び音声復号化装置の全体構成を示すブロック図である。この図において図1と同一の部分については同一の符号を付し、説明を省略する。

【0088】図4において、図1と比べて新たな構成は、過去フレームにおいて求められた適応音源符号から入力音声のピッチ周期を推定するピッチ推定手段40、42、推定されたピッチ周期から位相振幅特性の次数を決定する位相振幅特性次数決定手段41、43である。

【0089】以下、上記の追加された構成を中心に動作を説明する。

【0090】まず符号化部1において、ピッチ推定手段40は過去フレームにおいて求められた適応音源符号から現フレームの入力音声のピッチ周期を推定し、位相振幅特性次数決定手段41に出力する。適応音源符号は音声のピッチ周期に関わる情報を符号化したものであり、ピッチ周期の推定は、例えば過去フレームで求められた適応音源符号の頻度分布から出現頻度が大きい適応音源符号を求め、その適応音源符号に対応するピッチ周期を推定結果とする。

【0091】位相振幅特性次数決定手段41は、前記ピッチ周期に対応して位相振幅特性の次数を決定し、位相振幅特性分析手段32、位相振幅特性符号化手段33、位相振幅特性付加フィルタ34に出力する。前記位相振幅特性次数決定手段41においてピッチ周期に対する位相振幅特性の次数は、例えばピッチ周期が2～6msの時は40次、6～12msの時は80次、12ms以上では128次とする。

【0092】位相振幅特性分析手段32は、入力音声5と線形予測パラメータ補間手段10より入力される線形予測パラメータを用いて線形予測残差信号を生成し、フレーム毎に前記線形予測残差信号を前記位相振幅特性次数決定手段41より入力される次数で短期の位相振幅特性を分析して求め、位相振幅特性符号化手段33に出力する。位相振幅特性符号化手段33は前記次数に対応して、例えばベクトル量子化する際の量子化符号帳を前記次数により切替えるなどして、前記位相振幅特性を量子化し、それに対応する符号を多重化手段3に出力するとともに、量子化した位相振幅特性を遅延器25と位相振幅特性補間手段26に出力する。

【0093】遅延器25は前記量子化した位相振幅特性を1フレーム分時間遅延して位相振幅特性補間手段26に出力する。すなわち遅延器25は前フレームの量子化した位相振幅特性を位相振幅特性補間手段26に出力する。位相振幅特性補間手段26は前記遅延器25から入

力される前フレームの量子化した位相振幅特性と前記位相振幅特性符号化手段33から入力される現フレームの量子化した位相振幅特性を図10に示す方法で、例えば1/4フレームのサブフレーム周期毎に補間し、補間した位相振幅特性を位相振幅特性付加フィルタ34に出力する。

【0094】位相振幅特性付加フィルタ34は例えばFIRフィルタであり、前記位相振幅特性次数決定手段41から入力される次数に対応してフィルタ次数を変更し、適応音源符号帳12から出力されるサブフレーム長の適応音源ベクトル及び駆動音源符号帳13から出力されるサブフレーム長の駆動音源ベクトルにそれぞれ音源利得 $\beta$ 、 $\gamma$ を増幅器50、51で乗じ、これ等を加算器52で加算した音源信号に対して、前記補間した位相振幅特性を付加し、合成フィルタ11に出力する。合成フィルタ11は、線形予測パラメータ補間手段10より入力される補間した線形予測パラメータと前記位相振幅特性を付加した音源信号とを用いてサブフレーム長の合成音声を生成する。

【0095】最適音源探索手段14は、前記合成音声と入力音声5との該サブフレーム間における減算器53で算出した誤差信号の聴覚重み付き歪みを評価し、前記歪みが最小になる適応音源符号 $L$ 、駆動音源符号 $I$ 、音源利得 $\beta$ 、 $\gamma$ を求め、適応音源符号 $L$ と駆動音源符号 $I$ とを多重化手段3に出力するとともに、音源利得 $\beta$ 、 $\gamma$ を音源利得符号化手段15に出力する。音源利得符号化手段15は、前記音源利得 $\beta$ 、 $\gamma$ を量子化し、その符号を多重化手段3に出力する。

【0096】以上符号化が終了した後、多重化手段3は前記量子化した線形予測パラメータに対応する符号、量子化した位相振幅特性に対応する符号、適応音源符号 $L$ 、駆動音源符号 $I$ 、及び量子化した音源利得 $\beta$ 、 $\gamma$ に対応する符号を伝送路に送出する。

【0097】以上が実施の形態3に係る音声符号化装置の概要である。

【0098】次に、音声復号化装置の動作について説明する。

【0099】まず、多重化手段3の出力を受けた分離手段4は、その出力を分離し、  
適応音源符号 $L$ →適応音源符号帳16、ピッチ推定手段42

駆動音源符号 $I$ →駆動音源符号帳17

音源利得の符号→音源利得復号化手段18

位相振幅特性の符号→位相振幅特性復号化手段36

線形予測パラメータの符号→線形予測パラメータ復号化手段19

にそれぞれ出力する。

【0100】ピッチ推定手段42は、符号化側のピッチ推定手段40と同様の動作により過去フレームの適応音源符号から現フレームのピッチ周期を推定し、位相振幅

特性次数決定手段43に出力する。位相振幅特性次数決定手段43は、符号化側の位相振幅特性次数決定手段41と同様の動作により前記ピッチ周期に対応して位相振幅特性の次数を決定し、位相振幅特性復号化手段36と位相振幅特性付加フィルタ37に出力する。位相振幅特性復号化手段36は、前記位相振幅特性の次数に対応して、例えばベクトル量子化符号帳を前記次数により切替えるなどして、前記位相振幅特性の符号に対応する位相振幅特性を復号化し、遅延器29と位相振幅特性補間手段30に出力する。遅延器29は前記復号化した位相振幅特性を1フレーム分遅延して位相振幅特性補間手段30に出力する。すなわち遅延器29は前フレームの復号化した位相振幅特性を位相振幅特性補間手段30に出力する。位相振幅特性補間手段30は前記遅延器29から入力される前フレームの復号化した位相振幅特性と前記位相振幅特性復号化手段28から入力される現フレームの復号化した位相振幅特性を図10に示す符号化部1の位相振幅特性補間手段26と同様に、例えば1/4フレームのサブフレーム周期毎に補間し、補間した位相振幅特性を位相振幅特性付加フィルタ37に出力する。

【0101】位相振幅特性付加フィルタ37は、前記位相振幅特性の次数に対応してフィルタ次数を変更し、適応音源符号帳16から出力される適応音源ベクトル及び駆動音源符号帳17から出力される駆動音源ベクトルにそれぞれ音源利得復号化手段18から出力される音源利得 $\beta$ 、 $\gamma$ を増幅器54、55で乗じ、これ等を加算器56で加算して得られる音源信号に対して、前記補間した位相振幅特性を付加し、合成フィルタ22に出力する。合成フィルタ22は、線形予測パラメータ補間手段21より入力される線形予測パラメータと前記位相振幅特性を付加した音源信号とを用いて出力音声6を合成し、出力する。

【0102】以上が実施の形態3に係る音声復号化装置の概要である。

【0103】この実施の形態3によれば、入力音声のピッチ周期を適応音源符号から推定し、このピッチ周期に応じて位相振幅特性付加フィルタの次数を変更することにより、位相振幅特性の次数の情報を伝送する必要がないために少ない伝送情報量で音源信号の再現性を良好にし、合成音声の品質を向上させることができる。

#### 【0104】実施の形態4

つづいてこの発明の音声符号化装置の実施の形態4を図面に従って説明する。

【0105】図5は、この実施の形態4の音声符号化装置の構成を示すブロック図である。この図において、図9と同一の部分については同一の符号を付し、説明を省略する。

【0106】この実施の形態において新たな構成は、サブフレーム長より長い区間まで合成音声を生成する合成フィルタ44、サブフレーム長より長い区間において合

成音声の歪みを評価する最適音源探索手段45である。

【0107】以下、上記の新たな構成を中心に動作を説明する。

【0108】合成フィルタ44は、適応音源符号帳12から出力されるサブフレーム長の適応音源ベクトル及び駆動音源符号帳13から出力されるサブフレーム長の駆動音源ベクトルにそれぞれ音源利得 $\beta$ 、 $\gamma$ を増幅器50、51で乗じ、これ等を加算器53で加算した音源信号と、線形予測パラメータ補間手段10から入力される補間した線形予測パラメータとを用いてサブフレーム長+Nサンプルの合成音声を生成する。図6に、サブフレームと合成音声との時間関係を示す。

【0109】最適音源探索手段45は、前記合成音声と入力音声5との該サブフレーム+Nサンプルの区間における減算器53で算出した誤差信号の聴覚重み付き歪みを評価し、前記歪みが最小になる適応音源符号L、駆動音源符号I、音源利得 $\beta$ 、 $\gamma$ を求め、適応音源符号Lと駆動音源符号Iを多重化手段3に出力するとともに、音源利得 $\beta$ 、 $\gamma$ を音源利得符号化手段15に出力する。音源利得符号化手段15は、前記音源利得 $\beta$ 、 $\gamma$ を量子化し、その符号を多重化手段3に出力する。

【0110】以上符号化が終了した後、多重化手段3は前記量子化した線形予測パラメータに対応する符号、適応音源符号L、駆動音源符号I、及び量子化した音源利得 $\beta$ 、 $\gamma$ に対応する符号を伝送路に送出する。

【0111】以上がこの実施の形態の音声符号化装置に特徴的な動作である。

【0112】この実施の形態によれば、音源符号帳より最適な音源信号を選択する際に、サブフレーム長より長い区間の合成音声を生成し、入力音声との歪みを評価することにより、該サブフレームの音源信号から生成される合成音声の該サブフレーム以後に与える影響も加味して音源信号を決定するため、合成音声の再現性を良好にし、品質を向上させることができる。

【0113】実施の形態5. 上記実施の形態4では、従来の符号駆動線形予測符号化装置を基に合成音声の歪み評価をサブフレーム長+Nサンプルで行うものであるが、従来の位相振幅特性を符号化する符号駆動線形予測符号化装置を基に、サブフレーム長の音源信号を位相振幅特性付加フィルタに通しサブフレーム長+Nサンプルの位相振幅特性を付加した音源信号を生成し、これを合成フィルタに通しサブフレーム長+Nサンプルの合成音声を生成し、これを歪み評価に用いるものとしてもよい。

#### 【0114】実施の形態6

上記実施の形態4及び5では、サブフレーム長の音源信号からサブフレーム長+Nサンプルの合成音声を生成して歪み評価に用いているが、ピッチ周期などの情報を用いて音源信号をサブフレーム長+Nサンプルまで予測生成し、これからサブフレーム長+Nサンプルの合成音声を生成して歪み評価に用いるものとしてもよい。

【0115】実施の形態7. つづいてこの発明のデジタルフィルタ補間法の実施の形態を図面に従って説明する。

【0116】図7は、この実施の形態のデジタルフィルタ補間法を用いた音声復号化装置の構成を示すブロック図である。この図において、図11と同一の部分については同一の符号を付し、説明を省略する。

【0117】この実施の形態において新たな構成は、位相振幅付加フィルタと合成フィルタの補間特性を補償するフィルタ特性を求める補間補償特性算出手段46、合成音声に対して補間補償を行なう補間補償フィルタ47である。

【0118】以下、上記の新たな構成を中心に動作を説明する。

【0119】位相振幅特性復号化手段28は、位相振幅特性の符号に対応する位相振幅特性を復号化し、遅延器29、位相振幅特性補間手段30及び補間補償特性算出手段46に出力する。遅延器29は前記復号化した位相振幅特性を1フレーム分遅延して位相振幅特性補間手段30と補間補償特性算出手段46に出力する。位相振幅特性補間手段30は前記遅延器29から入力される前フレームの復号化した位相振幅特性と前記位相振幅特性復号化手段28から入力される現フレームの復号化した位相振幅特性を図10に示す方法で、例えば1/4フレーム周期毎に補間し、補間した位相振幅特性を位相振幅特性付加フィルタ31と補間補償特性算出手段46に出力する。

【0120】線形予測パラメータ復号化手段19は、線形予測パラメータの符号に対応する線形予測パラメータを復号化し、遅延器20、線形予測パラメータ補間手段21及び補間補償特性算出手段46に出力する。遅延器20は前記復号化した線形予測パラメータを1フレーム分時間遅延して線形予測パラメータ補間手段21と補間補償特性算出手段46に出力する。線形予測パラメータ

補間手段21は前記遅延器20から入力される前フレームの復号化した線形予測パラメータと前記線形予測パラメータ復号化手段19から入力される現フレームの復号化した線形予測パラメータを図10に示す方法で、例えば1/4フレーム周期毎に補間し、補間した線形予測パラメータを合成フィルタ22と補間補償特性算出手段46に出力する。

【0121】補間補償特性算出手段46は、前記前フレームの復号化した位相振幅特性と前記前フレームの復号化した線形予測パラメータから、前記位相振幅特性付加フィルタ31と合成フィルタ22を複合した1つのフィルタとみた場合の前フレームのフィルタ特性を求め、また、前記現フレームの復号化した位相振幅特性と前記現フレームの復号化した線形予測パラメータから前記位相振幅特性付加フィルタ31と合成フィルタ22を複合した1つのフィルタとみた場合の現フレームのフィルタ特性を求め、前記前フレームのフィルタ特性と前記現フレームのフィルタ特性を図10に示す方法で、例えば1/4フレーム周期毎に補間する。次に、前記補間した位相振幅特性の逆特性と前記補間した線形予測パラメータから合成フィルタの逆特性を求め、これら2つの逆特性を前記補間したフィルタ特性に乗算することにより補間補償特性を求め、補間補償フィルタ47に出力する。

【0122】ここで、フレームnで復号化された合成フィルタ22及び位相振幅特性付加フィルタ31の伝達関数は例えばそれぞれ式(1)、(3)で表現され、補間補償特性算出手段46はこれら2つのフィルタを複合した伝達関数を例えば式(5)のようにFIRフィルタで近似してフィルタ係数 $c_n(k)$ を求め、これを例えば式(6)に従ってフレームn-1とフレームn間で線形に補間してフィルタ係数 $c_i(k)$ を求め、これと式(2)、(4)から式(7)にしたがって補間補償フィルタの伝達特性 $X_i(z)$ を求める。

【0123】

【数5】

$$C_n(z) = P_n(z)H_n(z) = \sum_{k=0}^{N_c-1} c_n(k) \quad (5)$$

【0124】

【数6】

$$c_i(k) = w_i c_{n-1}(k) + (1 - w_i) c_n(k) \quad (k = 0, \dots, N_c - 1) \quad (6)$$

【0125】

【数7】

$$X_i(z) = \frac{1 + \sum_{k=1}^{N_h} a_i(k) z^{-k}}{\sum_{k=0}^{N_p-1} h_i(k) z^{-k}} \sum_{k=0}^{N_c-1} c_i(k) \quad (7)$$

【0126】補間補償フィルタ47は、合成フィルタ22から入力される合成音声に対して前記補間補償特性を与え、出力音声6を出力する。

【0127】以上がこの実施の形態のデジタルフィルタ補間法を用いた音声復号化装置に特徴的な動作である。

【0128】この実施の形態によれば、位相振幅特性付加フィルタと合成フィルタを別個に補間した際の補間特性を補償して、これら2つのフィルタを複合した1つのフィルタとしてみた場合の補間特性をとるように補間補償フィルタにより合成音声を補正することにより、合成音声の再現性を良好にし、その品質を向上させることができる。

【0129】なお、この実施の形態は音声復号化装置について述べているが、本デジタルフィルタ補間法は、音声符号化装置、音声復号化装置の他、複数のデジタルフィルタを構成要素に持つ全てのデジタル信号処理装置に適用することができる。

【0130】実施の形態8  
つづいてこの発明のデジタルフィルタ補間法の別の実施の形態を図面に従って説明する。

【0131】図8は、この実施の形態のデジタルフィルタ補間法を用いた音声復号化装置の構成を示すブロック図である。この図において、図11と同一の部分については同一の符号を付し、説明を省略する。

【0132】この実施の形態において新たな構成は、位相振幅特性付加フィルタと合成フィルタとを複合した一つのフィルタとしてみた場合の補間特性が得られるように位相振幅付加フィルタの特性を求め、その結果により位相振幅付加フィルタの特性を設定する特性補間設定手段としての位相振幅特性補間手段48である。

【0133】以下、上記の新たな構成を中心に動作を説明する。

【0134】位相振幅特性復号化手段28は、位相振幅特性の符号に対応する位相振幅特性を復号化し、遅延器29と位相振幅特性補間手段48に出力する。遅延器29は前記復号化した位相振幅特性を1フレーム分遅延して位相振幅特性補間手段48に出力する。

【0135】線形予測パラメータ復号化手段19は、線形予測パラメータの符号に対応する線形予測パラメータ

を復号化し、遅延器20、線形予測パラメータ補間手段21及び位相振幅特性補間手段48に出力する。遅延器20は前記復号化した線形予測パラメータを1フレーム分時間遅延して線形予測パラメータ補間手段21と位相振幅特性補間手段48に出力する。線形予測パラメータ補間手段21は前記遅延器20から入力される前フレームの復号化した線形予測パラメータと前記線形予測パラメータ復号化手段19から入力される現フレームの復号化した線形予測パラメータを図10に示す方法で、例えば1/4フレーム周期毎に補間し、補間した線形予測パラメータを合成フィルタ22に出力する。

【0136】位相振幅特性補間手段48は、前記前フレームの復号化した位相振幅特性と前記前フレームの復号化した線形予測パラメータから、前記位相振幅特性付加フィルタ31と合成フィルタ22を複合した1つのフィルタとみた場合の前フレームのフィルタ特性を求め、また、前記現フレームの復号化した位相振幅特性と前記現フレームの復号化した線形予測パラメータから前記位相振幅特性付加フィルタ31と合成フィルタ22を複合した1つのフィルタとみた場合の現フレームのフィルタ特性を求め、前記前フレームのフィルタ特性と前記現フレームのフィルタ特性を図10に示す方法で、例えば1/4フレーム周期毎に補間する。次に、前記補間した線形予測パラメータから合成フィルタの逆特性を求め、この逆特性を前記補間したフィルタ特性に乗算することにより補間補償をした位相振幅特性を求め、位相振幅特性付加フィルタ31に出力する。

【0137】ここで、フレームnで復号化された合成フィルタ22及び位相振幅特性付加フィルタ31の伝達関数は例えばそれぞれ式(1)、(3)で表現され、位相振幅特性補間手段48はこれら2つのフィルタを複合した伝達関数を例えば式(5)のようにFIRフィルタで近似してフィルタ係数 $c_n(k)$ を求め、これを例えば式(6)に従ってフレームn-1とフレームn間で線形に補間してフィルタ係数 $c_j(k)$ を求め、これと式(2)から式(8)にしたがって補間補償をした位相振幅特性を求め、これを位相振幅特性付加フィルタ31で実現するために $N_p$ 次のFIRフィルタで近似してフィルタ係数を求める。

【0138】

【数8】

$$P_i(z) = (1 + \sum_{k=1}^{N_h} a_i(k)z^{-k}) \sum_{k=0}^{N_e-1} c_i(k) \quad (8)$$

【0139】以上がこの実施の形態のデジタルフィルタ補間法を用いた音声復号化装置に特徴的な動作である。

【0140】この実施の形態によれば、位相振幅特性付加フィルタと合成フィルタを別個に補間した際の補間特性を補償して、これら2つのフィルタを複合した1つのフィルタとして見た場合の補間特性をとるように位相振

幅特性付加フィルタの特性を求めることにより、新たにフィルタを追加する必要がなく、既存のフィルタのみを用いても合成音声の再現性を良好にし、その品質を向上させることができる。

【0141】なお、この実施の形態は音声復号化装置について述べているが、本デジタルフィルタ補間法は、音声符号化装置、音声復号化装置の他、複数のデジ

タルフィルタを構成要素に持つ全てのデジタル信号処理装置に適用することができる。

【0142】実施の形態9. 上記実施の形態7及び8では複合したフィルタの補間特性を式(6)のようにFIRフィルタのフィルタ係数を線形に補間して求めており、これは該復号したフィルタのインパルス応答を時間波形上で補間していることになるが、インパルス応答の時間波形を例えばFFTして周波数軸上の表現に変換し、振幅特性と位相特性に分離し、これらを別々に補間するとしてもよい。

【0143】

【発明の効果】以上詳述したように、請求項1及び請求項4に記載の発明によれば、音声符号化装置で音源信号に短期の位相振幅特性を付加するフィルタの次数を可変とし、また音声復号化装置でも符号化装置に対応するようにフィルタの次数を可変としたので、音源信号の位相振幅特性の再現性が良くなるようにフィルタ次数を選択して符号化、復号化でき、高品質の音声を合成することができる。

【0144】また、請求項2に記載の発明によれば、音源信号に位相振幅特性を付加するフィルタの次数を入力音声の特徴に応じて変更するようにしたので、入力音声に対応して音源信号の位相振幅特性の再現性が良くなるフィルタ次数を求められ、高品質の音声を合成することができる。

【0145】また、請求項3及び請求項5に記載の発明によれば、音源信号に位相振幅特性を付加するフィルタの次数を音声符号に応じて変更するようにしたので、伝送情報量を増加させることなく入力音声に対応して音源信号の位相振幅特性の再現性が良くなるフィルタ次数を求められ、高品質の音声を合成することができる。

【0146】また、請求項6に記載の発明によれば、音声符号化側で音源信号に位相振幅特性を付加するフィルタの次数を可変とし、また音声復号化側でも符号化側に対応するようにフィルタの次数を可変としたので、音源信号の位相振幅特性の再現性が良くなるようにフィルタ次数を選択して符号化、復号化でき、高品質の音声を合成することができる。

【0147】また、請求項7に記載の発明によれば、音源信号情報を符号化する際に、該区間以後において生成される合成音声を予測して生成し、入力音声と合成音声との歪みを評価するようにしたので、後続する区間の合成音声に悪影響を与えない音源信号を生成でき、高品質の音声を合成することができる。

【0148】また、請求項8に記載の発明によれば、複数のデジタルフィルタから構成される複合デジタルフィルタで、フィルタ特性を補間する際に、複合デジタルフィルタのフィルタ特性を補間した特性が得られるように、フィルタ特性を補償するフィルタを挿入するようにしたので、デジタルフィルタの補間特性を向上

させることができる。

【0149】また、請求項9に記載の発明によれば、音声符号化装置または音声復号化装置に請求項8に記載の複合デジタルフィルタを用いるようにしたので、合成音声の再現性を良好にし、その品質を向上させることができる。

【0150】また、請求項10に記載の発明によれば、複数のデジタルフィルタから構成される複合デジタルフィルタで、フィルタ特性を補間する際に、特性補間設定手段により複合デジタルフィルタのフィルタ特性を補間する特性を演算し、この演算結果により、各フィルタの特性を設定するようにしたので、デジタルフィルタの補間特性を向上させることができる。

【0151】また、請求項11に記載の発明によれば、音声符号化装置または音声復号化装置に請求項10に記載のデジタルフィルタ補間法を用いるようにしたので、合成音声の再現性を良好にし、その品質を向上させることができる。

【図面の簡単な説明】

【図1】この発明の実施の形態1の構成を示すブロック図である。

【図2】この発明の実施の形態1における位相振幅特性符号化手段の構成の一例を示すブロック図である。

【図3】この発明の実施の形態2の構成を示すブロック図である。

【図4】この発明の実施の形態3の構成を示すブロック図である。

【図5】この発明の実施の形態4の構成を示すブロック図である。

【図6】この発明の実施の形態4における音源探索における音源信号、合成音声の時間関係を示す説明図である。

【図7】この発明の実施の形態7の構成を示すブロック図である。

【図8】この発明の実施の形態8の構成を示すブロック図である。

【図9】従来の符号駆動線形予測符号化復号化装置の一例の構成を示すブロック図である。

【図10】従来の符号化フレーム構成の一例を示す説明図である。

【図11】従来の音源信号の位相振幅特性を符号化する符号駆動線形予測符号化復号化装置の一例の構成を示すブロック図である。

【図12】音源信号の位相振幅特性を符号化する符号駆動線形予測符号化復号化装置における音源信号から合成音声を生成する過程を示す説明図である。

【図13】ピッチ周期と位相振幅特性付加フィルタのインパルス応答長の関係を示す説明図である。

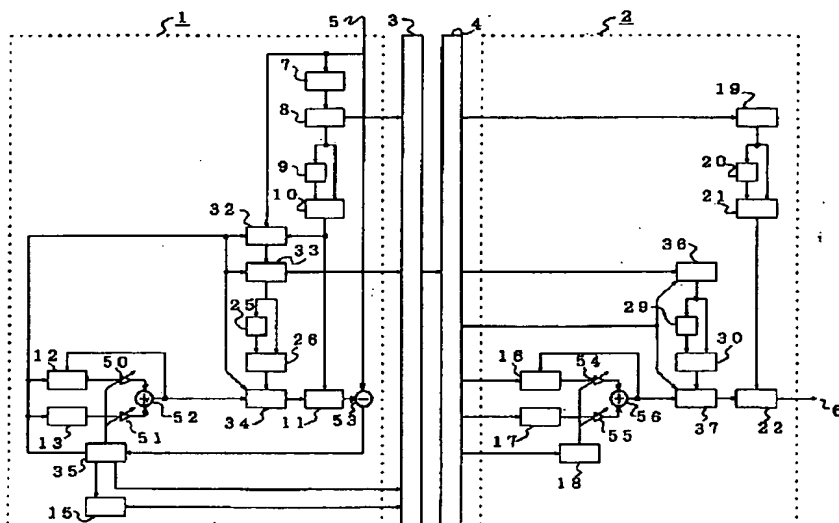
【符号の説明】

1：符号化部、

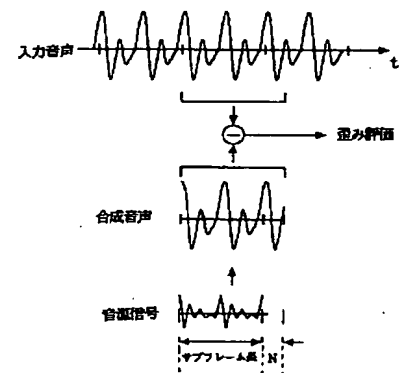
2：復号化部

- 3 : 多重化手段、  
 5 : 入力音声、  
 7 : 線形予測パラメータ分析手段、  
 9、20 : 遅延器、  
 11、22 : 合成フィルタ、  
 13、17 : 駆動音源符号帳、  
 15 : 音源利得符号化手段、  
 19 : 線形予測パラメータ復号化手段  
 23 : 位相振幅特性分析手段、  
 25、29 : 遅延器、  
 27、31 : 位相振幅特性付加フィルタ
- 4 : 分離手段  
 6 : 出力音声  
 8 : 線形予測パラメータ符号化手段  
 10、21 : 線形予測パラメータ補間手段  
 12、16 : 適応音源符号帳  
 14 : 最適音源探索手段  
 18 : 音源利得復号化手段
- 28 : 位相振幅特性復号化手段、  
 33 : 位相振幅特性符号化手段  
 34、37 : 位相振幅特性付加フィルタ  
 35 : 最適音源探索手段、  
 38 : ピッチ抽出手段、  
 40、42 : ピッチ推定手段  
 41、43 : 位相振幅特性次数決定手段  
 44 : 合成フィルタ、  
 46 : 補間補償特性算出手段、  
 48 : 位相振幅特性補間手段
- 32 : 位相振幅特性分析手段  
 36 : 位相振幅特性復号化手段  
 39 : 位相振幅特性次数決定手段  
 45 : 最適音源探索手段  
 47 : 補間補償フィルタ  
 50、51、54、  
 55 : 増幅器  
 52、56 : 加算器  
 53 : 減算器

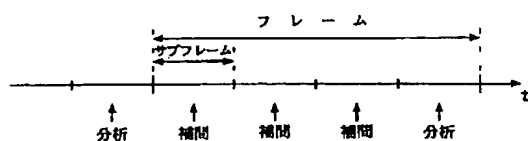
【図1】



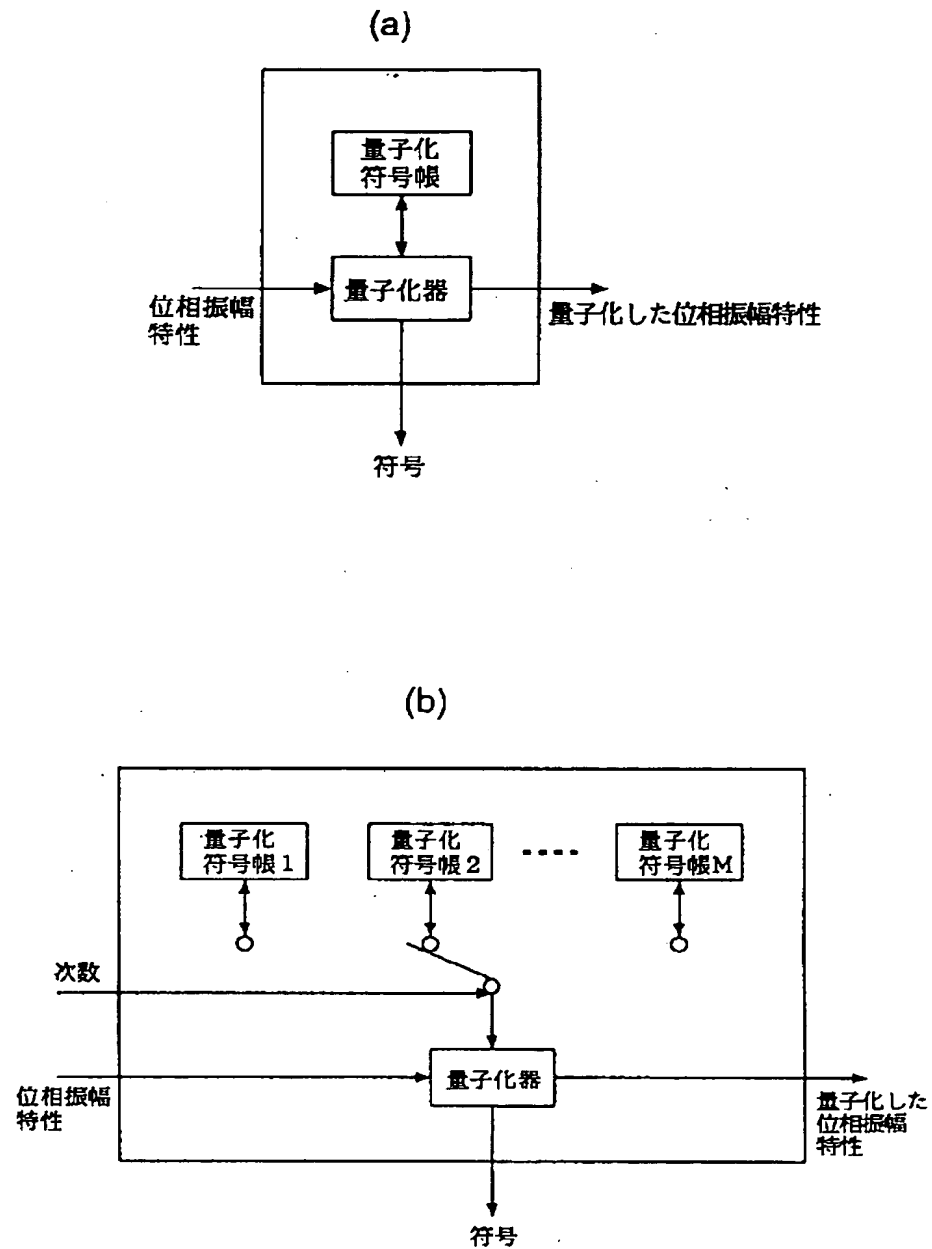
【図6】



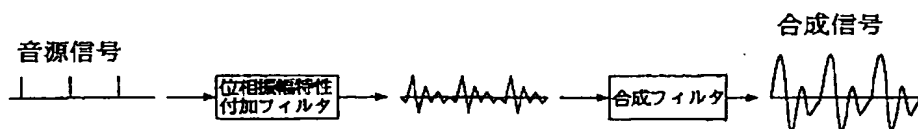
【図10】



【図2】

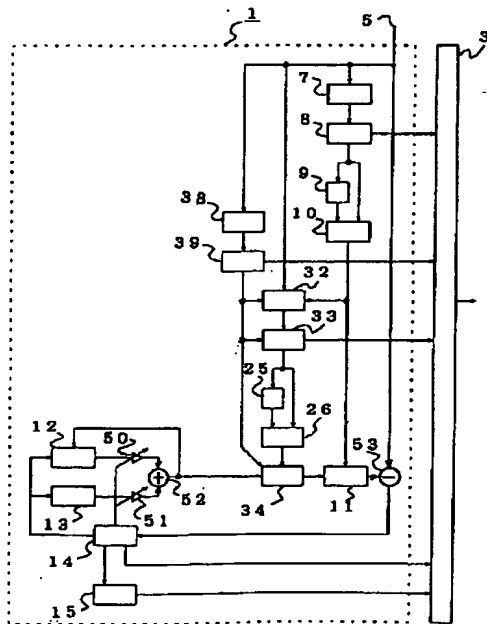


【図12】

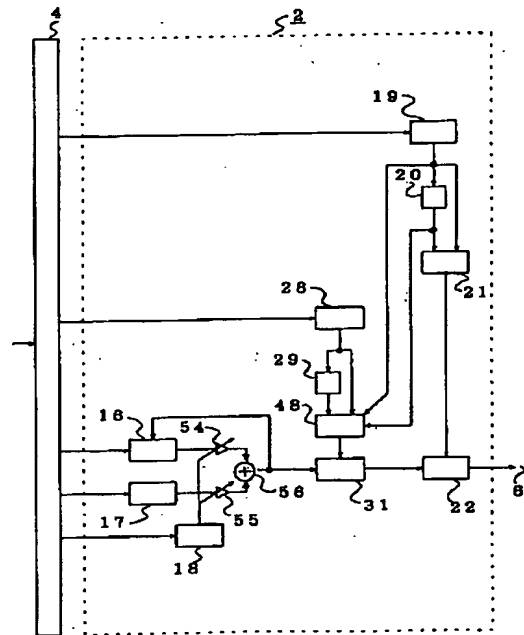




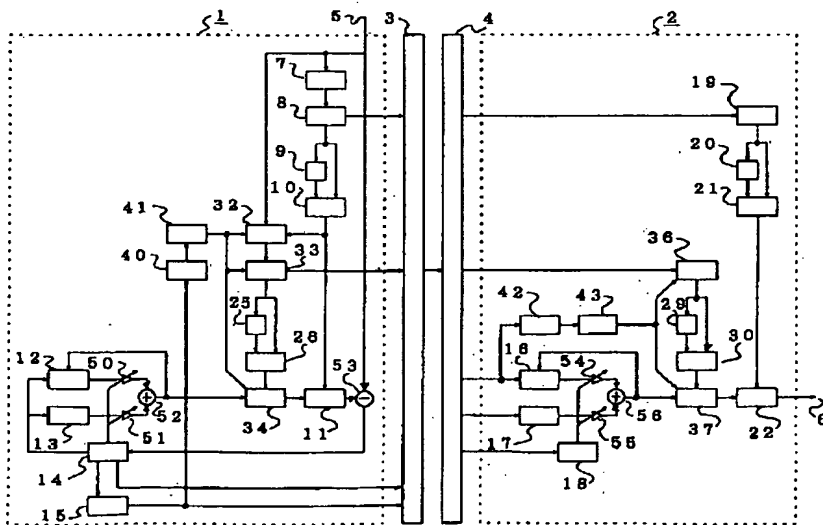
【図3】



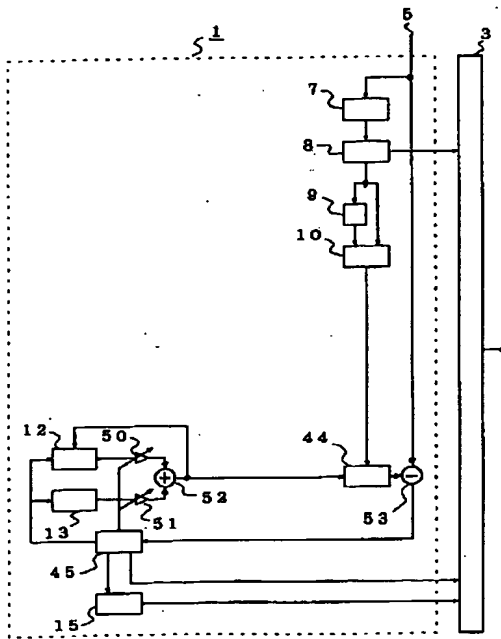
【図8】



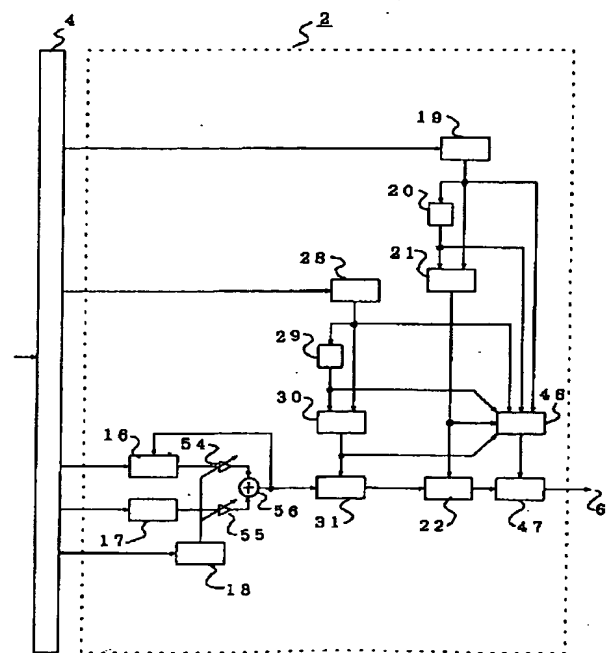
【図4】



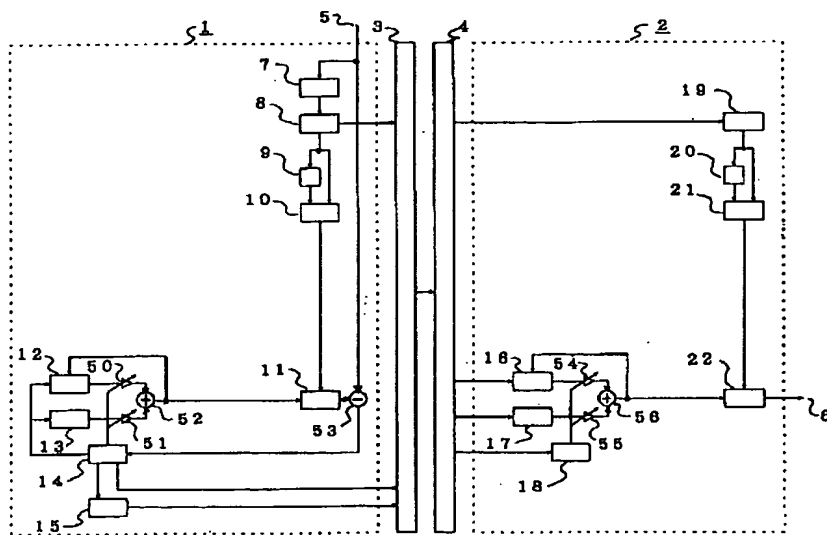
【図5】



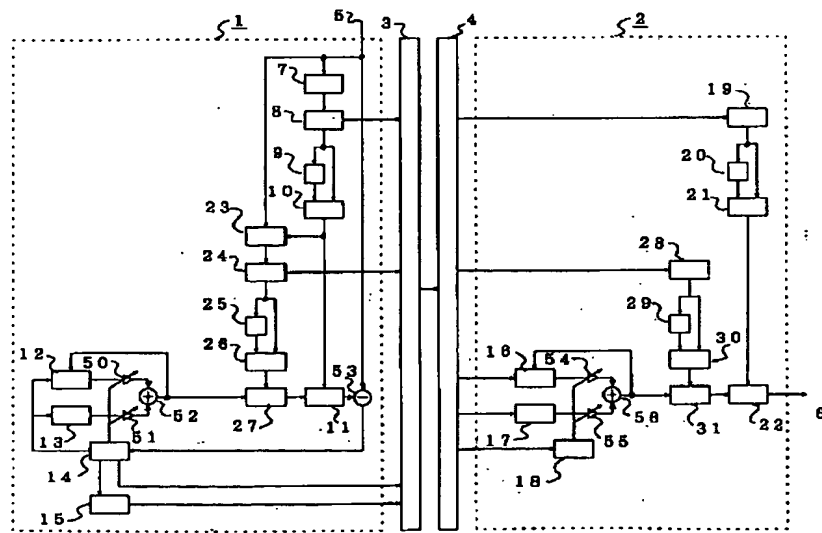
【図7】



【図9】



【図11】



【図13】

